Towards mapping the late quaternary vegetation change of Europe

Thomas Giesecke, Basil Davis, Simon Brewer, Walter Finsinger, Steffen Wolters, Maarten Blaauw, Jacques-Louis de Beaulieu, K.D. Bennett, Heather Binney , Ralph M. Fyfe, Marie-José Gaillard, Graciela Gil-Romera, W.O. van der Knaap, Petr Kunes , Norbert Kühl, Jacqueline F.N. van Leeuwen, Michelle Leydet, André F. Lotter, Elena Ortu, Malte Semmler, Richard H. W. Bradshaw

Abstract

Over the last 30 years the number of absolutely dated pollen diagrams has increased and many have been submitted to the European Pollen Database (EPD). This allows the construction of more precise maps of Holocene vegetation change across Europe. Site chronologies and taxonomic considerations are prerequisites for the construction of maps. From the beginning of palaeoecological research chronological control improved through time. Until recently chronological information in the EPD was expressed in uncalibrated radiocarbon years and most chronologies are to date based on this time scale. Here we present new chronologies for most of the datasets stored in the EPD based on a calibrated scale. These are general purpose chronologies fit for most continental scale questions, but include some chronological information derived from vegetation history that needs to be considered in different applications. The steps taken to generate the new dataset are described and the rational for a new classification system for age uncertainties is introduced. Taxonomic particularities of the data stored in the EPD are explained. An example is given of how the database can be queried to select samples with appropriate age control as well as the suitable taxonomic level to answer a specific research question.

Keywords: EPD, European Pollen Database, Chronology, Age-Depth Relationships, Age uncertainty, *Plantago lanceolata*

Introduction

Like all spatial features, plant distributions and their changes through time are best appreciated when displayed on a map. One of the first maps illustrating the distributional change of a species during the Holocene was the reconstruction of the maximal range of *Corylus avellana* in Sweden (Andersson, 1909). Anderson collected the reports of hazel nut finds without knowing their stratigraphic position or having any means of dating them. Peat stratigraphy provided the chronological framework for von Post’s first pollen diagrams, where he used the distinct shift in peat humification around 2700 cal. BP to align the diagrams (von Post 1918). He realised that general changes in forest composition through time could be used to assign the strata to distinct time periods. Changes in peat stratigraphy had previously been used to divide the Holocene into periods with distinct wet or dry climate (the Blytt–Sernander scheme; Birks & Seppä, 2010) and von Post used these periods and characterized their general forest composition. Using this chronological framework and pollen analysis from about 250 profiles, von Post (1924) published the first maps depicting changes in tree abundance through time and space. Dividing the Holocene into six periods he constructed maps over southern Sweden for different species, where abundance is represented by circle diameter. Szafer (1935) developed the mapped representation of pollen analytical results further as he treated pollen percentages from 152 sites across Poland like elevation data and produced the first isopollen maps. The later recognition of the different herbaceous pollen types and, in particular, the identification of cereal pollen by Firbas (1937) meant that archaeology and history could be used to provide chronological information on the vegetation history over the last few thousand years. When Firbas (1949) presented isopollen maps for central Europe he restricted himself to four broad periods that follow the Blytt–Sernander scheme.

With the advent of radiocarbon dating in the 1950’s the possibilities of comparing pollen diagrams changed completely. It was now possible to plot pollen diagrams against an absolute time scale and compare the timing of events rather than their relative position with respect to some other common pattern. Initially radiocarbon age determinations were expensive and pollen stratigraphical zones were a generally accepted dating scheme in many parts of Europe. Thus radiocarbon dating was used to derive ages for the regional pollen zone boundaries and individual features of pollen diagrams. As the expansion of *Picea abies* is one of the major changes in Finnish pollen diagrams this feature was often directly dated with the radiocarbon method. Thus Aario (1965) and Aartolahti (1966) were able to construct isochrone maps depicting the timing when *P. abies* populations expanded in Finland. In eastern North America, Davis (1976) used a selection of pollen diagrams for which radiocarbon based chronologies could be established and interpreted the time at which trees arrived at the different sites. By interpolating between the arrival times for the different species she created a set of isochrone maps depicting their postglacial spread.

The publication of “An Atlas of Past and Present Pollen Maps for Europe” by Huntley and Birks (1983) allowed for the first time to study postglacial changes in the abundance of 46 pollen taxa through time and space on a sub-continental scale. The isopollen maps assembled in this atlas were mainly compiled from printed pollen diagrams, with a total of 423 sites throughout Europe. Two thirds of these diagrams had some independent dating control that in total amounts to more than 1500 radiocarbon age determinations. However, compiling data from printed diagrams has limitations and it was realised that databases containing the pollen counts were needed for different applications including climate reconstructions (Fyfe et al., 2009). In consequence the European Pollen Database (EPD) was created in 1992 alongside with the North American Pollen Database

(NAPD). Unlike the EPD, the NAPD had a predecessor, the COHMAP pollen 'database' managed by T. Webb III, which was the basis for sub-continental climate reconstructions and data model comparisons, while it also yielded maps of past vegetation change (e.g. Prentice et al., 1991). This was followed by a comprehensive set of isopollen and derived biome maps for the last 21,000 years covering much of North America and mainly using the data stored in the NAPD (Williams 2004). The maps were constructed as 1000 year time slices using the calendar age scale and collecting pollen samples in 500 year wide windows. However, the age models for the underlying individual pollen diagrams were based on the radiocarbon time scale. Also the age models stored alongside the pollen data in the EPD are until now mostly based on the uncalibrated radiocarbon time and the few maps of individual taxa that are based on the EPD were presented using this time scale (e.g. Brewer et al. 2002, Magri et al. 2006). While this time-scale is adequate to document the diachronous nature of the spread of taxa, the offset between ^{14}C and calendar ages biases interpretations of the dynamics of these abundance changes and hampers the comparison to other proxy information with calendar time scales.

After almost 100 years of palynological research in Europe the general vegetation history is well known, while updated or detailed maps only exist for selected species like *Picea* (Giesecke and Bennett 2004, Latalowa, van der Knaap 2006) or for particular regions or countries like Poland (Ralska-Jasiewiczowa et al. 2004), Spain (Carrión et al. 2010) and the Alps (van der Knaap et al. 2005). At the same time there is a renewed interest in aspects of vegetation history from neighbouring disciplines. Genetic markers are being used to reveal ancestries between tree populations across Europe, thus indicating routes of postglacial spread of trees (Petit et al. 2002, Magri et al. 2006). Climate simulations for past millennia make it possible to explore past vegetation distributions with vegetation models. These allows evaluating the accuracy of climate simulations as well as the vegetation models and eventually give the opportunity to explore the relationship between shifting tree distributions and climate change over the last 20,000 years (Giesecke et al. 2007, Pearman et al. 2008). The constructions of maps of past vegetation, as well as most other research questions utilizing pollen data, require that the samples can be assigned to an accurate age estimate. Studies that need accurate sample-age estimates include the reconstruction of continental scale trends in climate change (Davis et al. 2003) or the reconstruction of landscape openness through time (Gaillard et al. 2010, Nielsen et al. 2012).

The growing need for calendar time scale chronologies for the pollen data in the EPD is being met with the work outlined in this manuscript. We also present a system indicating the uncertainty on the age estimate for individual samples and discuss constraints and considerations for estimating the ages and uncertainties. We describe how the uncertainty information can be used for sample selection. Finally the chronologies were used to map the distribution of selected pollen types for the last 15,000 years based on the data in the EPD and taxonomic considerations are discussed. All work presented here is based on the version of the database released the 19th of March 2012.

Dates – building blocks of chronologies

Here a date is understood as a particular depth in the stratigraphy with some information on the age the material was deposited. This may be an absolute age estimate obtained from the decay of a radioactive isotope (e.g. radiocarbon) or an event that is found in several records and can thus be used for correlation of two or more records. The best relative marker horizon is the fallout of a volcanic ash (tephra), where it is spread over a large area and easily identifiable. Also clear signals in proxy records that have a regional or continental cause and are thus synchronous can be used as a date. However, where changes in the proxy are used for chronological information these changes become unavailable for analysis as this would lead to circular arguments. For example the onset of the Holocene is a climatic event of global scale that can be identified in many pollen diagrams. By using this feature as a date it is afterwards not possible to investigate whether the palynologically detectable onset of the Holocene was synchronous across Europe. When building chronologies for individual sites these relative dates have to be expressed as absolute ages and estimates of their uncertainty provided. Absolute age estimates, like radiocarbon dates, are reported with an uncertainty, but the sample dated may not have been ideal and the resulting age estimate may be erroneous. While additional age information may be included when building chronologies, some absolute ages may be omitted and the resulting mix of dates are here referred to as control points. Apart from annually laminated sediments or well dated and easily identifiable tephras, the most common control points for Holocene and late-Glacial pollen diagrams are based on radiocarbon measurements. However, the coverage of well-dated diagrams is still uneven and it is therefore necessary to include other sources of age estimates. We reviewed and evaluated all different control points in the EPD, omitted some, which could not be confirmed and added others like the onset of the Holocene where the chronologies would otherwise be unconfined. The labelling of the control points was refined to identify the different characters and uncertainties associated to the age estimates (Table 1), and to provide more fine-grained information for subsequent investigations.

Additional site information that can help to constrain the possible chronologies includes knowledge on basin formation and development, detailed regional vegetation and archaeological information, lithology and additional proxy records. This full set of constraining chronological information was rarely available when constructing the here presented chronologies. However, we used general knowledge, regional vegetation history and checked original publications as often as possible to evaluate chronologies and control points.

Control points and their uncertainty

Radiometric dates

Radiocarbon

The most common control points are radiocarbon dates and the version of the EPD used here stores currently 5080 such dates. However, not all of them are used in the construction of chronologies, as authors often reported all ^{14}C dates from the sequences even if some were in their opinion inaccurate. Radiocarbon dates that appear to be too old or too young can be caused by a variety of reasons. The best known of these is perhaps the uptake of old carbon by aquatic plants or mosses (Björck & Wohlfarth 2001). Hard water and reservoir effects particularly affect bulk sediment dates, which used to be common as conventional radiocarbon dating required large amounts of carbon containing material (Grimm et al. 2009). The accelerator mass spectrometry (AMS) technique works with much smaller quantities making it possible to work with terrestrial macrofossils in many situations. In theory, plant fragments, seeds or fruits will be embedded in sediments shortly after their tissue assimilated atmospheric carbon, thus yielding a most accurate date. But using small samples also means that a tiny amount of contamination is sufficient to affect a date (Wohlfarth et al. 1998). In a search for sites that are dated with a high degree of confidence, Blois et al. (2011) classify radiocarbon dates based on their chance to be inaccurate. We did not apply a classification to the dates but used this type of information when evaluating them based on their stratigraphical context and excluded spurious dates. Thus we used the reported uncertainty of the date regardless whether it came from a conventional bulk date or an AMS date from terrestrial macrofossils.

Lead 210

Some cores held in the EPD were dated in their uppermost part using the radioactive isotope of lead (^{210}PB). This isotope has a half-life of 22.3 years and is thus useful to obtain a time control for the last 150 years. However unlike radiocarbon its production varies regionally, requiring more complex models to obtain sample ages (Appleby 2001). These age models are often constrained by the abundance of other radioactive isotopes such as Caesium (^{137}Cs) which show a peak for the time of peak atomic weapon testing and, in many European regions, the Chernobyl fallout. We did not evaluate the provided age estimates and included them in the chronologies with the reported uncertainty.

Table 1: List of control points and associated age and uncertainty

Type of control point	Age in cal. BP	uncertainty	error distribution	R_code
Radiocarbon		from date	full calibrated probability distribution	C14
Lead 210		From dating	Gaussian	PB2
Core top- known	year of samplin g	± 25	half-Gaussian	TOP_A
Core-top-known, not sampling year	- 40	± 50	half-Gaussian	TOP_B

Core top- estimated	0	± 250	half-Gaussian	TOP_C
Annual laminations		1% per year	Gaussian	ANL
Tephra (e.g. Laarcher See)	age of deposition (12880)	of age determination (± 120)	Gaussian	TEF (LST)
Local pollen biostratigraphic events		± 150	Gaussian	POL_A
Regional pollen biostratigraphic: clearly identifiable and well constrained		± 250	Gaussian	POL_B
Regional pollen biostratigraphic: vague in character or regional occurrence		± 500	Gaussian	POL_C
Beginning of the Holocene	11500	± 250	Gaussian	HOL
Onset Younger Dryas	12800	± 500	Gaussian	YDO
Onset Bølling	14700	± 500	Gaussian	BOL
Estimated bottom age		± 500	Gaussian	BOT

Lithological dates

Core tops

Most late Quaternary cores were collected in situations where sediment is still accumulating. In these situations the sediment surface can be assigned to the year of sampling, which adds an important control point for the chronologies. Depending on the coring device it is not always possible to obtain the youngest sample, which is a particular problem in lakes. Wherever possible the relevant publications were examined to decide whether the core top is modern and how it was collected. Each core top was classified based on the information available to us at the time and slightly different ages are assigned to the different classed dates (Table 1). In situations where no information was available to us, the pollen composition of the topmost samples was evaluated and if it suggested modern vegetation composition a category C was assigned with an age of 1950 (0 BP) and an uncertainty of 250 years. Uncertainty in time was used here also to indicate the probability of loss of material as in this case the topmost sample would be slightly older. As future ages are impossible for sediments that were deposited in the recent past we used a half Gaussian error distribution to allow uncertainty in the past, but not into the future.

Annual Laminations

Lakes with continuously annual laminated sediments provide the best possible age control and the EPD stores a few pollen diagrams from such sites. These datasets have ages assigned to all samples and thus each sample is therefore a control point. However, varves are rarely perfect leading to differences in varve-counting replication and slumps created gaps of unknown length creating uncertainties in sample ages (e.g. Brauer et al. 2000). In addition, annually laminated sediments don't always occur all the way to the sediment water interface and the floating varve chronologies then have to be dated by radiocarbon or linked to tephra layers. Unless error estimates were provided with the chronology we assigned an error of 1% of the ascribed age.

Tephra layers

Important control points for many European pollen diagrams are tephra layers that can be identified and assigned to an age. Northern Europe has mainly received volcanic ash from the eruptions of Icelandic volcanos and eruptions from various volcanos in the Mediterranean provide chronological information in this region (Davies et al., 2012). The most common tephra in central European cores is the Laacher See Tephra (LST), which is easily recognised and dated to 12880 ± 120 cal. BP (Brauer et al. 1999). Recent developments in tephrochronology enabling the extraction of small amounts of shards (microtephra), allow for the development of a spatially more extended tephrostratigraphy over Europe (e.g . Lane et al., 2012).

Biostratigraphic dates

A large proportion of pollen grains deposited at anyone site are derived from the regional vegetation (Sugita 1993). Therefore two pollen diagrams from adjacent basins will reflect similar changes in regional vegetation composition and if such features are well dated for one diagram their age can be used as control points in the second diagram. Regional vegetation change itself was often determined or triggered by changes in climate and human land use which shows some synchrony over large European regions (Giesecke et al. 2011). The onset of the Holocene climate warming has caused vegetation shifts globally and late-Glacial fluctuations like the Younger Dryas can be identified in many regions worldwide. The age of these dramatic climate shifts is being determined with increasing accuracy (Rasmussen et al. 2006) and where the pollen analytical correspondence is identified this can serve as a valuable control point. While Rasmussen et al. (2006) dated the onset of the Holocene in Greenland to 11,700 cal. BP, well dated lake records in Europe suggest a somewhat younger age for the palynological signal (e.g. Blockley et al. 2008) and we therefore assumed an age of 11500 ± 250 cal. BP. The onset of the Holocene (HOL) was used as an additional control point in many diagrams while the onset of the Bølling (BOL) and the Younger Dryas (YDO) were less frequently used. The rise of *Corylus avellana* and *Alnus* is broadly parallel in some European regions (Giesecke et al. 2011) and was used as a control point with a 500 year uncertainty in these regions. The timing of the decline in *Ulmus* pollen varies somewhat between regions but its regional age was also used as a palynological control point if no independent dating was available. Where authors had used more palynological control points in their data submission to the EPD we tried to assess how the age information was derived and assigned uncertainties or omitted control points that seemed vague. Knowledge on the timing of archaeological periods can also provide dating control, where the pollen diagram is indicating individual settlement phases. Also the introduction of cultivated plants like *Juglans*, may serve as a date, but this is potentially only a “younger than” date as it is uncertain if the first cultivation or later expansions are captured. In some cases historical events like the 30-years war in central Europe (Giesecke 2001) can be identified and provide very accurate and precise time control.

The addition of each palynological control point reduces the degrees of freedom of the overall dataset and palynological control points were therefore only included where the chronologies could otherwise not sufficiently constrained without them. Some pollen diagrams stored in the EPD are derived from landscapes that changed little over the course of the Holocene or contain features that have not yet been well dated in other diagrams from the region. In these situations, we did not attempt to introduce control points other than the onset of the Holocene and the resulting chronologies are poorly constrained.

Estimated bottom ages

In recently glaciated regions the time of glacier retreat is often known with some confidence. Thus lakes or bogs in these regions can usually not have started to accumulate sediments before the glaciers retreated. However, particularly in lakes, initial inorganic sedimentation was often rapid and slowed with the establishment of dense vegetation cover. Down-core extrapolations from younger dates often cannot capture this initial rapid sediment accumulation and where constraining knowledge is available this is valuable in establishing the chronologies for the lowermost sections.

Building chronologies and propagating uncertainties

The choice of method

With the exception of annually laminated sediments, a standard profile has fewer control points than samples and thus there is a need to make inferences about the sample age based on the available age-depth information. An increasing number of methods and programs are becoming available to construct age-depth relationships and propagate the uncertainties of the age determinations to the samples (Bennett, 1994; Heegaard et al. 2005, Ramsey 2008). Although newly available Bayesian approaches (e.g. Blaauw and Christen, 2011) are powerful tools and provide more robust uncertainties for the age estimates of the samples, these techniques need much supervision and computing power and were therefore inadequate for this task, given the number and diversity of sites. Instead we choose to work with simple linear interpolation and smoothing spline regression as implemented in CLAM (Blaauw, 2010). These two different techniques represent inherently different philosophies in the construction of age-depth relationships. Linear interpolations force the age-depth relationship through the centre of each control point assuming that these points are the only places where accurate knowledge is available. In consequence control points causing age-depth reversals have to be resolved as negative accumulation rates are impossible. Smoothing splines are in effect locally weighted regressions, fitting a cubic spline to the overall set of control points, with a smoothing term that avoids the need for exact fitting to each control point. Unless the factor controlling the amount of smoothing is effectively removed, the resulting age-depth relationship does not necessarily go through every control point. Thus small reversals in the control points do not need to be eliminated manually. Fitting smoothing splines in CLAM has the additional advantage that it uses the full power of the uncertainties associated to the control points and where vegetation history is used as a control point with a large uncertainty it works like a guide rather than a tight constrain. This reduces some of the circularity that is introduced when using vegetation history to date vegetation history, as the final age assigned to a vegetation change may differ from the age assumed for this event.

Implementation

All age-depth relationships were constructed using the R code CLAM version 1.0.2 (Blaauw, 2010), which offered the advantages that radiocarbon dates are internally calibrated so that the dates can be stored as uncalibrated dates together with control points expressed as calendar year ages. After evaluating all control points, they were run as a batch yielding two age-depth relationships based on linear interpolation and smoothing spline fitting with a default smoothing factor of 0.3. The two alternative models were visually evaluated, selecting the more appropriate one, with a preference for the smoothing spline. However, smoothing splines had sometime problems where the density of dates varied markedly within a sequence, and could not be produced for sites with a low number (<4) of control points. Where none of the two age-depth relationships provided an adequate description, the control points were edited and if necessary the smoothing factor changed. Uncertainties were obtained in CLAM through drawing random samples from the probability distributions of each control point and fitting 1000 age-depth relationships to these points. The resulting uncertainties may be non-symmetrical, especially for the smoothing spline based age models. The sample age is obtained as the highest probability age based on the distribution of estimated ages from the 1000 runs and the uncertainties are provided as 95% confidence intervals. The half Gaussian uncertainty distribution assigned to the core top was only considered where smoothing splines were fitted. In linear interpolations, using a half Gaussian uncertainty for the top would have resulted in uncertainties biased toward lower ages, due to the constraint of exactly fitting control points. The resulting uncertainties for core tops range between the past and future, in particular, where the linear model was chosen.

Table 2: Classification of sample age uncertainty

Maximum distance to the nearest date in years	Stars
2000	1
1000	2
500	3
straight segment	+1

An additional uncertainty classification

The propagation of the age uncertainty from the control points to the samples is satisfactory where a high density of control points is available. However, linear interpolation and smoothing splines do this in different ways. For linear interpolations, sample uncertainties decrease between control points, as they are effectively constrained by two points. In smoothing splines, these depend more on the curvature of the age depth and are often larger between control points. Where the time interval between control points is large none of the two procedures can adequately describe the uncertainty on the ascribed sample age. Therefore we created an uncertainty classification that reflects mainly the density of control points to complement the assignment of uncertainty through inter and extrapolation. The classification is additive and samples are assigned to the lowest class (a single star) where the estimated sample age is within 2000 years of the nearest control point (Table 2). Additional stars are given at 1000 and 500 year proximity to the nearest control point. In addition to these three stars, that characterize proximity to the nearest control point, a further star is given to samples that are situated in a straight section of the sequence. This straightness star describes the necessity of additional control points to adequately describe complex sequences with large changes in sediment accumulation rates or hiatuses. The straightness star is given to a sample where within the nearest 4 control points the sediment accumulation rate is constant or changes less than 20%. Thus only sequences with at least four control points can obtain a straightness star. The evaluation is based on the position of the sample relative to the control points and independent of the interpolation procedure.

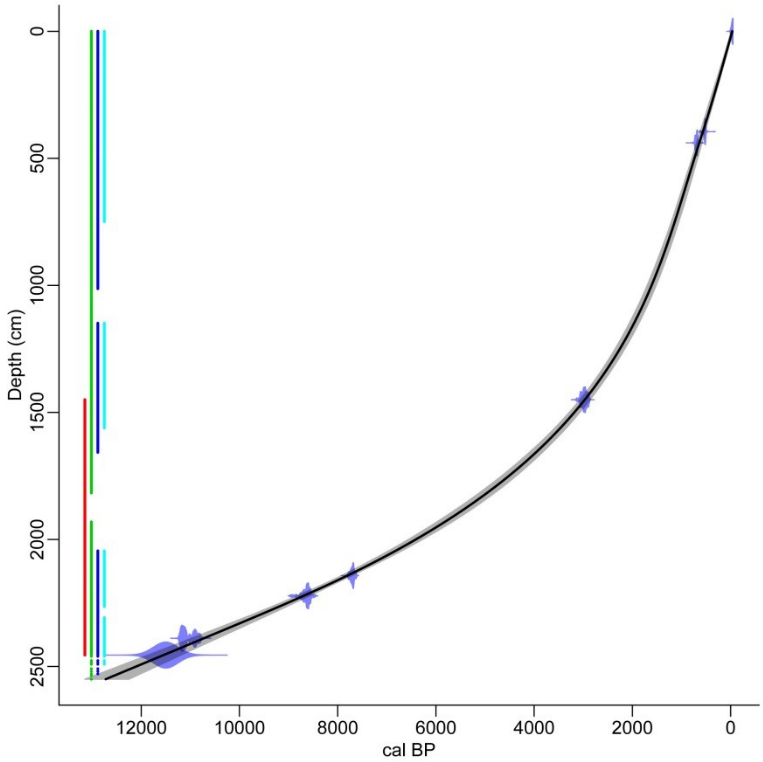


Fig.1 Age-depth relationship for Treppelsee, exemplifying the application of the classification based on control point density and sequence complexity indicated on the left side beside the depth scale. The first (red) bar marks samples placed within a segment of the core bracketed by at least four control points within which the sediment accumulation changes less than 20%. The three following bars (green, dark blue, light blue) reflect the proximity of a sample to the nearest control points. The lowermost control point represents the palynologically determined onset of the Holocene set to 11500 cal. BP with an associated standard deviation of 250 years (this can be identified from its obvious Gaussian distribution). The addition of this biostratigraphical control point constrains the extrapolation which would otherwise yield too old ages. The blue polygons represent the calibrated age range as a distribution, where the height of the polygon represents the probability of that age.

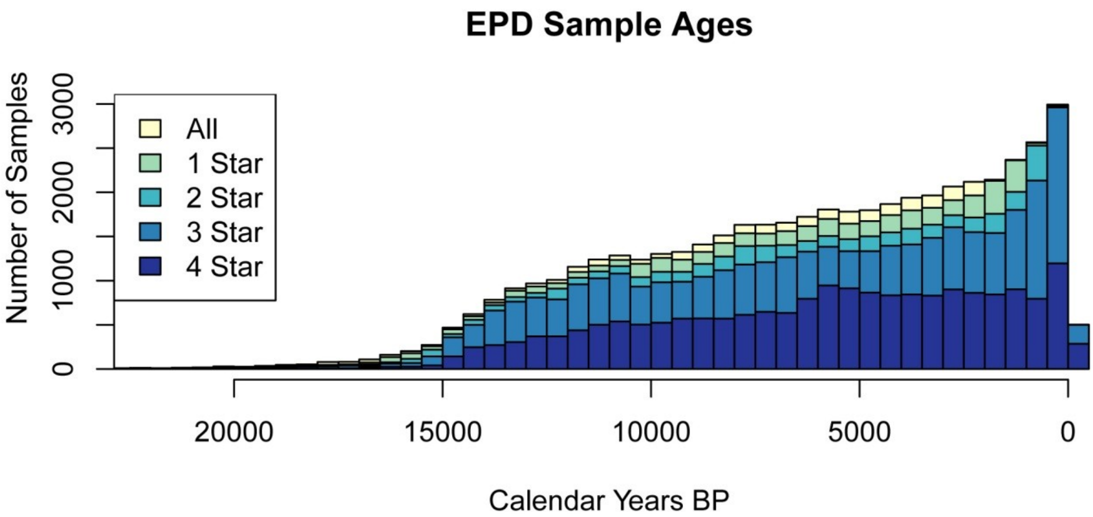
The age-depth relationship inferred for the sediment core from Treppelsee (Giesecke 2001) gives a good example, illustrating how the star classification helps describing uncertainties for sample ages (Fig 1). The more than 25m long core was dated with 6 radiocarbon dates, two of which are situated close to each other. The sediment water interface was obtained using a freeze corer giving a high confidence to the assignment of the core top to the year of coring. The beginning of the Holocene could be inferred from the pollen diagram and is used as an additional control point with a standard deviation of 250 years. From the control point marking the beginning of the Holocene until a radiocarbon date at 3000 cal. BP the increase in sedimentation rate stays below 20% giving this section an additional star. Thus samples with inferred ages around 5500 years cal. BP have no control point within 2000 years, but their inferred ages are constrained by the general age-depth relationship, giving them a single star. The sediment accumulation rate increased between 7000 and 3000 years cal. BP, while it is uncertain where exactly this increases started and whether it was gradual or sudden. The confidence limits assigned to the samples through curve fitting do not reflect this uncertainty in this section of the core, but the decreasing number of stars helps to reflect the uncertainty of this change. Early Holocene samples attained 3 to 4 stars due to the better density of control points and here the confidence limits for the sample ages describe the uncertainties derived from the control points in an adequate way. Also the most recent samples in the profile are well constrained by control points, here, the addition of a single control point would fulfil the straight segment constraint.

Limiting chronologies

The resulting chronologies were assessed though checking of individual pollen diagrams plotted against sample age as well as interpolations of species abundances for different time periods across Europe. These evaluations made it necessary to restrict the analyses to a section of the pollen diagram, which was implemented as a restriction of the chronology. Restrictions were set for mainly two different reasons: i) In some cases it was impossible to constrain the age depth model towards the bottom or top of the core and the extrapolations yielded unrealistic age estimates. For example, late-Glacial diagrams from areas that were not glaciated or where deglaciation occurred much before the onset of limnic sedimentation could often not be constrained and extrapolating ages over several ten thousand years seemed unrealistic in the absence of any constraint. Restrictions towards the top are less common but were set in cases where the core-top was apparently not modern, its age could not be otherwise estimated and extrapolation yielded inappropriate ages. ii) Reworked pollen grains that are out of their biostratigraphic context are common towards the bottom of cores and during the late Glacial, where erosion cut into older organic deposits. On a site specific basis these reworked pollen grains are usually easily identified, while their assessment in collections of samples from particular time slices is difficult and leads to misinterpretations. Although, some of these samples may be adequately assigned to the age of their deposition, the pollen composition of these samples does not reflect the vegetation near the sampling site at that time. These cut-off limits were determined as sediment depth and stored in a separate table.

General purpose chronologies and sample selection

The chronologies that were produced following the above outlined rational and procedure will rarely be the best chronology that may be achievable for any given site. Where pollen data was used to derive control points the chronologies are not entirely independent from assumed timing of vegetation change. Nevertheless these chronologies provide more independent age control for the mapping of vegetation change than was possible in the dataset compiled by Huntley and Birks (1983). The chronologies should be adequate for most continental scale questions and database queries. We like to think of them as general purpose chronologies and stress that users should consider whether they are adequate to peruse their research questions, particularly if these are site-specific.



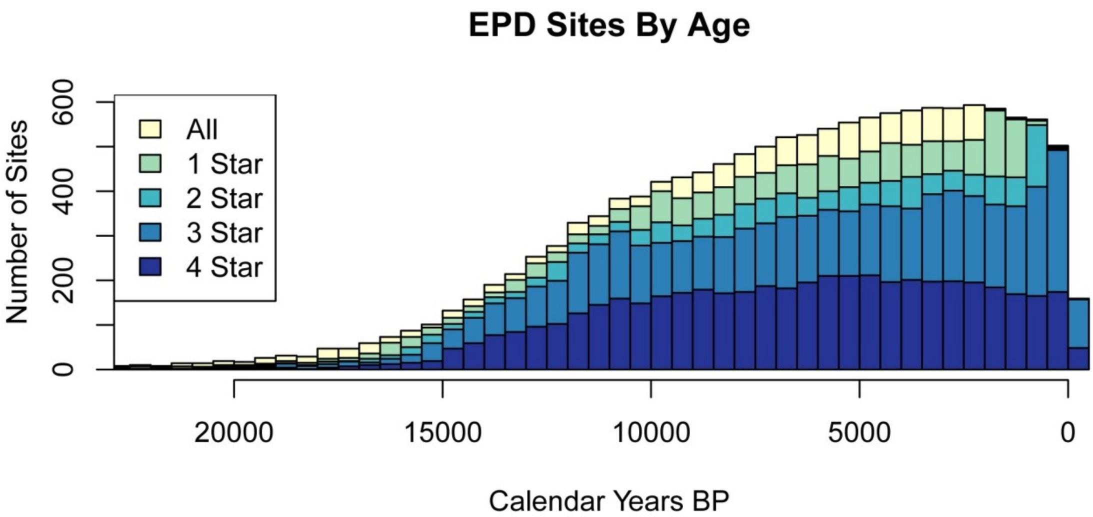


Fig. 2: Histograms depicting the distribution of samples and the number of cores available in the EPD over the last 25000 years in 500 year bins. The different colours illustrate the proportion of samples or samples in core segments that were classified with 3 and 4 stars indicating a good control point density for most samples.

While the new calendar time scale chronologies are an important step towards the mapping of European vegetation change and related questions, the computation of sample age uncertainties on a database wide scale provides new opportunities for sample selection and database queries. Individual sites are often well dated for some sections, while only a vague age estimates can be provided for other samples. The twofold uncertainty classification provided here allows more flexibility on the inclusion of samples for a particular analysis. Where high confidence in age control is necessary, for example, comparisons of vegetation patterns and archaeological periods or to investigate vegetation responses to climate perturbations like the 8.2 event, the investigator can restrict the analysis to samples with three stars and uncertainties within 250 years. Other applications, like the mapping of late Quaternary vegetation change on a continental scale do perhaps not require highly accurate ages but need a good spatial spread of samples and here one star and a 1000 year uncertainty envelop may suffice. It has to be remembered that the star classification needs to be used in conjunction with the sample uncertainty derived through line fitting. For example a diagram from a region with marked vegetation changes with regionally known ages may have a satisfactory density of control points reaching two or more stars even if no or only few radiocarbon dates are available. However, the large uncertainties of these types of control points result in large confidence intervals for individual samples, exceeding confidence intervals of 1000 years. In constructing the chronologies we aimed to use few control points based on vegetation history and the vast majority of control points comes from radiocarbon dates. However, the effect of the onset of the Holocene as a control point is visible in the histograms depicting the number of samples or core segments assigned to 3 and 4 stars (Fig. 2).

Taxonomy – what are we mapping?

Since von Post's lecture in 1916 pollen identification has advanced substantially, leading to the creation of even more detailed pollen determination keys, pollen atlases (e.g. Fægri et al. 1989, Moore et al. 1991, Reille 1992, 1995, 1998, Punt et al. 1976–2009, Beug 2004), and pollen reference collections. Thus the detail of taxonomic identification has considerably improved through time. Unfortunately, nomenclature and definition of pollen types is not consistent. Depending on the determination key in use, different labs or schools use different nomenclatures, which make it difficult for database users to synthesise all relevant information. For example, many genera of the Asteroideae constitute a single pollen morphological type, but this type is named diversely e.g. *Solidago*-type in Fægri et al. (1989), *Aster*-type in Moore et al. (1991) and *Senecio*-type in Beug (2004). Similarly, an analyst working in northern Europe may decide to name a pollen type according to a particular species (e.g. *Sanguisorba minor*) while a colleague working in southern Europe has to deal with more species or genera that produce the same pollen type (e.g. *Sanguisorba* spp. and *Sarcopoterium*). Other nomenclatural obstacles arise when authors apply different plant nomenclatures (e. g. Umbelliferae/Apiaceae or *Fallopia*/*Bilderdykia*). Eventually, pollen type identification and hence naming also depends on the pollen preservation of a particular site as well as the investigator and her or his experience, which usually grows over time.

The EPD aims to harmonize those nomenclatural inconsistencies by implementation of a standard catalogue of common pollen type names and the introduction of a simple hierarchy. The EPD-nomenclature is geographically confined to Europe and plant names follow the Flora Europaea (Tutin et al. 1964–

1993). It finally enables database users to access the full data set of particular taxa regardless whether different pollen type names were originally assigned. However, the EPD stores the original taxonomy and nomenclature as supplied by the author. A table (P_VARS) within the database links the author's pollen type names to a common EPD nomenclature retaining all details carried by the original naming. The name of an EPD pollen type consists of either 1) the name of a plant taxon (e.g. *Calluna vulgaris*, *Fagus*, Cruciferae); 2) two plant taxon names separated by a slash (e.g. *Humulus/Cannabis*); 3) the plant taxon name with suffix -type (e.g. *Rumex acetosa*-type). In most cases the EPD nomenclature follows the naming of recognised determination keys or pollen atlases (e.g. Fægri et al. 1989, Moore et al. 1991, Reille 1992, 1995, 1998, NEPF 1976–2009, Beug 2004) with the reference of the type name given. In case a new EPD name is created, a full description or explanation will be provided.

The P_VARS table is also meant to link up each pollen taxon to a higher level collective taxon. For example the higher level collective taxon for *Secale*, *Zea mays*, *Hordeum*-type, *Triticum*-type and *Avena*-type is Cerealia-type, which also comprises all ambiguous cereal pollen determinations such as cf. *Triticum*, *Avena/Triticum*, Cerealia undiff., etc. Consequently database users can map either single well determined pollen taxa or the sum of all cereal pollen. An example from southern Europe is, experience and preservation provided, the distinction of *Quercus* into *Quercus robur*-type, *Quercus cerris*-type and *Q. ilex*-type. The latter types predominantly include evergreen *Quercus* species. This group with some shared ecological features is also an interesting level for mapping. However, because of the limited discriminability of *Quercus* pollen a map of identifications of evergreen *Quercus* pollen can only be a minimum map where the true distribution and abundance of these pollen types is greater than the mapped area and or abundance. If circumstances do not allow the recognition of *Quercus cerris*-type or *Q. ilex*-type such pollen would be identified only to the genus level *Quercus*. All distinguished *Quercus*-types and *Quercus* together constitute the higher level collective taxon *Quercus*. When mapping the higher level collective taxon *Quercus* database users receive a map of all *Quercus* pollen records including the more precise determinations.

Currently the P_VARS table provides some taxonomic hierarchy that makes it possible to collect pollen identifications at taxonomic and ecological meaningful levels. However, this hierarchy is not in all cases adequate or detailed enough and work is in progress to improve these links. For the current mapping the data stored in the EPD we considered the taxonomic groupings in the P_VARS table and made adjustments where necessary.

Applying the chronological information to the making of maps

The work outlined here had the overarching goal to produce maps depicting the data currently held in the EPD as a basis for further spatial analysis. While the resulting maps are presented elsewhere (Brewer et al. 2012), an example with a slightly different motivation is shown here to illustrate further considerations in the making of the maps.

One may be interested in assessing the strength of human land use in Europe during particular archaeological periods, such as the medieval forest clearance, Migration period, Iron Age, Bronze Age and the Neolithic period. *Plantago lanceolata* is a good indicator for human land-use and its presence and abundance may give some indication on the strength of land-use. To be careful of over interpretations we restrict the maps to identifications of *P. lanceolata* and *P. lanceolata*-type. Pollen diagrams differ in their sample resolution and some have a sample every 50, others every 250 years and it is therefore better to collect samples over a common period rather than at a particular time and we here choose a time window of 500 years. To make sure the samples considered really come from within this period we restrict the search to samples that were classified with at least 3 stars and an uncertainty of less than or equal to 250 years.

The resulting average values in the 500-year time window can be plotted on a map (Fig. 3) displaying the abundances using colour or symbol size and a classification scheme for easy visualization. Averaging pollen percentages over a time window leads to higher effective pollen sums and increased probability of detecting a grain. Thus the presence of pollen grains in the 500-year time window needs to be considered in a differentiated way, as the sole occurrence may bear different information from its presence at low abundance. The threshold of 0.1% was used here which represents the occurrence of a single grain in a count of 1000 pollen.

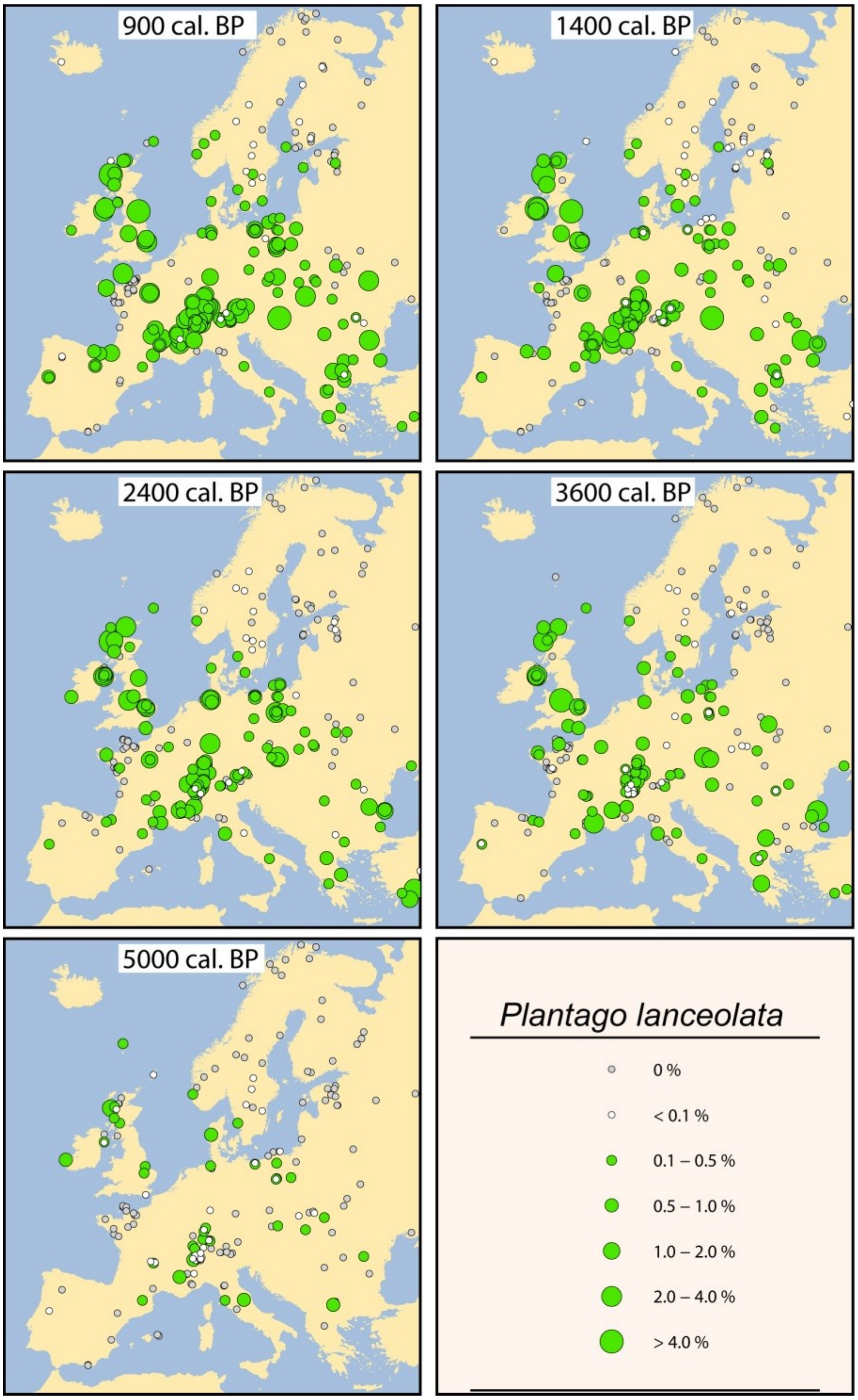


Fig. 3: Presence and abundance of *Plantago lanceolata* pollen during five time periods, chosen to represent the medieval forest clearance, Migration period, Iron Age, Bronze Age and a Neolithic period.

Concluding remarks

Here we present chronologies for the large majority of sites currently stored in the EPD, which are based on a calendar time scale and associated assessment of uncertainties for the inferred sample ages. A variety of control points were evaluated for 1036 sites and age-depth relationships constructed based on linear interpolation and preferably smoothing splines. These line fitting techniques were used to propagate the error from the uncertainties of the control points to the samples using the R-code CLAM. To complement these sample age confidences a classification capturing the dating density and the complexity of the age-depth relationship is provided. It thus becomes possible to select samples based on different requirements on the accuracy of the age control. This is providing a general purpose chronology fit for most continental scale questions. However, we encourage any user working with a small number or individual sites to review the individual chronologies and where necessary construct new ones. It also needs to be noted that we did not construct chronologies for all sites in the EPD, but worked only with sites where some chronological information had been submitted to the EPD or where previous efforts were made to establish a chronology.

Acknowledgments

We wish to thank all those who have contributed data to the EPD.

Supplementary data

The control points used as well as the derived chronologies, uncertainties and classifications are stored in Pangaea (doi link here).

References

Aario R (1965) Die Fichtenverhäufigung im Lichte von 14C Bestimmungen und die Altersverhältnisse der Finnischen Pollenzonen. Bulletin de la Commission geologique de Finlande 218:215–231

Aartolahti T (1966) Über die Einwanderung und die Verhäufigung der Fichte in Finnland. Annales Botanici Fennici 3:368–379

Appleby PG (2001) Chronostratigraphic techniques in recent sediments. In: Last WM, Smol JP (eds) Tracking Environmental Change Using Lake Sediments. Volume 1: Basin Analysis, Coring, and Chronological Techniques. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 171-203

Bennett KD (1994) Confidence intervals for age estimates and deposition times in late-Quaternary sediment sequences. Holocene 4:337-348

Beug HJ (2004) Leitfaden der Pollenbestimmung für Mitteleuropa und angrenzende Gebiete. Pfeil, München.

Birks HJB, Seppa H (2010) Late-Quaternary palaeoclimatic research in Fennoscandia - A historical review. Boreas 39 (4):655-673. doi:10.1111/j.1502-3885.2010.00160.x

Björck S, Wohlfarth B (2001) 14C chronostratigraphic techniques in paleolimnology. In: Last WM, Smol JP (eds) Tracking Environmental Change Using Lake Sediments. Volume 1: Basin Analysis, Coring, and Chronological Techniques. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 205-245

Blaauw M (2010) Methods and code for 'classical' age-modelling of radiocarbon sequences. Quaternary Geochronology 5 (5):512-518. doi:10.1016/j.quageo.2010.01.002

Blaauw M, Andres Christen J (2011) Flexible Paleoclimate Age-Depth Models Using an Autoregressive Gamma Process. Bayesian Analysis 6 (3):457-474. doi:10.1214/11-ba618

Blockley SPE, Ramsey CB, Lane CS, Lotter AF (2008) Improved age modelling approaches as exemplified by the revised chronology for the Central European varved lake Soppensee. Quaternary Science Reviews 27 (1-2):61-71. doi:10.1016/j.quascirev.2007.01.018

Blois JL, Williams JW, Grimm EC, Jackson ST, Graham RW (2011) A methodological framework for assessing and reducing temporal uncertainty in paleovegetation mapping from late-Quaternary pollen records. Quaternary Science Reviews 30 (15-16):1926-1939. doi:10.1016/j.quascirev.2011.04.017

Brauer A, Endres C, Gunter C, Litt T, Stebich M, Negendank JFW (1999) High resolution sediment and vegetation responses to Younger Dryas climate change in varved lake sediments from Meerfelder Maar, Germany. Quaternary Science Reviews 18 (3):321-329. doi:10.1016/s0277-3791(98)00084-5

Brauer A, Gunter C, Johnsen SJ, Negendank JFW (2000) Land-ice teleconnections of cold climatic periods during the last Glacial/Interglacial transition. Climate Dynamics 16 (2-3):229-239. doi:10.1007/s003820050016

Brewer S, Cheddadi R, de Beaulieu JL, Reille M, Data c (2002) The spread of deciduous Quercus throughout Europe since the last glacial period. Forest Ecology and Management 156 (1-3):27-48. doi:10.1016/s0378-1127(01)00646-6

Carrion JS, Fernandez S, Gonzalez-Samperiz P, Gil-Romera G, Badal E, Carrion-Marco Y, Lopez-Merino L, Lopez-Saez JA, Fierro E, Burjachs F (2010) Expected trends and surprises in the Lateglacial and Holocene vegetation history of the Iberian Peninsula and Balearic Islands. Review of Palaeobotany and Palynology 162 (3):458-475. doi:10.1016/j.revpalbo.2009.12.007

Davies SM, Abbott PM, Pearce NJG, Wastegard S, Blockley SPE (2012) Integrating the INTIMATE records using tephrochronology: rising to the challenge. Quaternary Science Reviews 36:11-27. doi:10.1016/j.quascirev.2011.04.005

Davis BAS, Brewer S, Stevenson AC, Guiot J (2003) The temperature of Europe during the Holocene reconstructed from pollen data. Quaternary Science Reviews 22 (15-17):1701-1716

Davis MB (1976) Pleistocene biogeography of temperate deciduous forests. Geoscience and Man 13:13-26

Fægri K, Iversen J, Kaland PE, Krzywinski K (1989) Textbook of Pollen Analysis. Wiley, Chichester.

Firbas F (1937) Der pollenanalytische Nachweis des Getreidebaus. Zeitschr f Bot 31:447-478

Firbas F (1949) Spät- und nacheiszeitliche Waldgeschichte Mitteleuropas nördlich der Alpen, Allgemeine Waldgeschichte. Gustav Fischer, Jena

Fyfe RM, de Beaulieu JL, Binney H, Bradshaw RHW, Brewer S, Le Flao A, Finsinger W, Gaillard MJ, Giesecke T, Gil-Romera G, Grimm EC, Huntley B, Kunes P, Kuhl N, Leydet M, Lotter AF, Tarasov PE, Tonkov S (2009) The European Pollen Database: past efforts and current activities. Vegetation History and Archaeobotany 18 (5):417-424. doi:10.1007/s00334-009-0215-9

Gaillard MJ, Sugita S, Mazier F, Trondman AK, Brostrom A, Hickler T, Kaplan JO, Kjellstrom E, Kokfelt U, Kunes P, Lemmen C, Miller P, Olofsson J, Poska A, Rundgren M, Smith B, Strandberg G, Fyfe R, Nielsen AB, Alenius T, Balakauskas L, Barnekow L, Birks HJB, Bjune A, Bjorkman L, Giesecke T, Hjelle K, Kalnina L, Kangur M, van der Knaap WO, Koff T, Lageras P, Latalowa M, Leydet M, Lechterbeck J, Lindbladh M, Odgaard B, Peglar S, Segerstrom U, von Stedingk H, Seppa H (2010) Holocene land-cover reconstructions for studies on land cover-climate feedbacks. Climate Of The Past 6 (4):483-499. doi:10.5194/cp-6-483-2010

Giesecke T (2001) Pollenanalytische und sedimentchemische Untersuchungen zur natürlichen und anthropogenen Geschichte im Schlaubetal. Sitzungsberichte der Gesellschaft Naturforschender Freunde zu Berlin 39:89-112

Giesecke T, Bennett KD (2004) The Holocene spread of Picea abies (L.) Karst. in Fennoscandia and adjacent areas. J Biogeogr 31 (9):1523-1548

Giesecke T, Bennett KD, Birks HJB, Bjune AE, Bozilova E, Feurdean A, Finsinger W, Froyd C, Pokorny P, Roesch M, Seppa H, Tonkov S, Valsecchi V, Wolters S (2011) The pace of Holocene vegetation change - testing for synchronous developments. Quaternary Science Reviews 30 (19-20):2805-2814. doi:10.1016/j.quascirev.2011.06.014

Giesecke T, Hickler T, Kunkel T, Sykes MT, Bradshaw RHW (2007) Towards an understanding of the Holocene distribution of Fagus sylvatica L. J Biogeogr 34 (1):118-131

Grimm EC, Maher LJ, Jr., Nelson DM (2009) The magnitude of error in conventional bulk-sediment radiocarbon dates from central North America. Quaternary Research 72 (2):301-308. doi:10.1016/j.yqres.2009.05.006

Heegaard E, Birks HJB, Telford RJ (2005) Relationships between calibrated ages and depth in stratigraphical sequences: an estimation procedure by mixed-effect regression. Holocene 15 (4):612-618

Huntley B, Birks HJB (1983) An atlas of past and present pollen maps for Europe 0-13000 years ago. Cambridge University Press, Cambridge

Lane CS, Blockley SPE, Lotter AF, Finsinger W, Filippi ML, Matthews IP (2012) A regional tephrostratigraphic framework for central and southern European climate archives during the Last Glacial to Interglacial transition: comparisons north and south of the Alps. Quaternary Science Reviews 36:50-58. doi:10.1016/j.quascirev.2010.10.015

Latalowa M, van der Knaap WO (2006) Late Quaternary expansion of Norway spruce Picea abies (L.) Karst. in Europe according to pollen data. Quaternary Science Reviews 25 (21-22):2780-2805

Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gomory D, Latalowa M, Litt T, Paule L, Roure JM, Tantau I, van der Knaap WO, Petit RJ, de Beaulieu JL (2006) A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. New Phytologist 171 (1):199-221

Moore PD, Webb JA, Collinson ME (1991) Pollen analysis. Blackwell Science, Oxford.

Nielsen AB, Giesecke T, Theuerkauf M, Feeser I, Behre K-E, Beug H-J, Chen S-H, Christiansen J, Dörfler W, Endtmann E, Jahns S, de Klerk P, Kühl N, Latałowa M, Odgaard BV, Rasmussen P, Stockholm JR, Voigt R, Wiethold J, Wolters S (2012) Quantitative reconstructions of changes in regional openness in north-central Europe reveal new insights into old questions. Quaternary Science Reviews 47 (0):131-149. doi:10.1016/j.quascirev.2012.05.011

Pearman PB, Randin CF, Broennimann O, Vittoz P, van der Knaap WO, Engler R, Le Lay G, Zimmermann NE, Guisan A (2008) Prediction of plant species distributions across six millennia. Ecology Letters 11 (4):357-369. doi:10.1111/j.1461-0248.2007.01150.x

Petit RJ, Brewer S, Bordacs S, Burg K, Cheddadi R, Coart E, Cottrell J, Csaikl UM, van Dam B, Deans JD, Espinel S, Fineschi S, Finkeldey R, Glaz I, Goicoechea PG, Jensen JS, Konig AO, Lowe AJ, Madsen SF, Matyas G, Munro RC, Popescu F, Slade D, Tabbener H, de Vries SGM, Ziegenhagen B, de Beaulieu JL, Kremer A (2002)

Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. Forest Ecology And Management 156 (1-3):49-74

Prentice IC, Bartlein PJ, Webb T (1991) Vegetation and climate change in eastern North-America since the Last Glacial Maximum. Ecology 72 (6):2038-2056

Punt W, Clarke GCS, Blackmore S, Hoen PP, Stafford PJ (eds) 1976–2009. The Northwest European Pollen Flora I–IX. Elsevier, Amsterdam

Ralska-Jasiewiczowa M, Latalowa M, Wasylikowa K, Tobolski K, Madeyska E, Wright HE, Charles Turner K (2004) Late Glacial and Holocene history of vegetation in Poland based on isopollen maps. W. Szafer Institute of Botany, Polish Academy of Sciences, Krakow, Poland

Ramsey CB (2008) Deposition models for chronological records. Quaternary Science Reviews 27 (1-2):42-60. doi:10.1016/j.quascirev.2007.01.019

Rasmussen SO, Andersen KK, Svensson AM, Steffensen JP, Vinther BM, Clausen HB, Siggaard-Andersen ML, Johnsen SJ, Larsen LB, Dahl-Jensen D, Bigler M, Rothlisberger R, Fischer H, Goto-Azuma K, Hansson ME, Ruth U (2006) A new Greenland ice core chronology for the last glacial termination. Journal of Geophysical Research-Atmospheres 111 (D6). doi:10.1029/2005jd006079

Reille M (1992) Pollen et spores d'Europe et d'Afrique du Nord. Laboratoire de Botanique historique et Palynologie, Marseille

Reille M (1995) Pollen et spores d'Europe et d'Afrique du Nord. Supplement 1. Laboratoire de Botanique historique et Palynologie, Marseille

Reille M (1998) Pollen et spores d'Europe et d'Afrique du Nord. Supplement 2. Laboratoire de Botanique historique et Palynologie, Marseille

Sugita S (1993) A Model of Pollen Source Area for an Entire Lake Surface. Quaternary Research 39 (2):239-244

Szafer W (1935) The significance of isopollen lines for the investigation of the geographic distribution of trees in the Post-Glacial period. Bulletin de l'Academie Polonaise des Sciences B:235–239

Tutin TG, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA (eds.) 1964–1993. Flora Europaea. University Press, Cambridge

van der Knaap WO, van Leeuwen JFN, Finsinger W, Gobet E, Pini R, Schweizer A, Valsecchi V, Ammann B (2005) Migration and population expansion of Abies, Fagus, Picea, and Quercus since 15000 years in and across the Alps, based on pollen-percentage threshold values. Quaternary Science Reviews 24 (5-6):645-680. doi:10.1016/j.quascirev.2004.06.013

von Post L (1818) Skogsträdpollen i sydsvenska torvmosselagerföljder. Forhandlinger ved de 16 Skandinaviske Naturforskeresmøte 1916:433–465 von Post L (1924) Ur de sydsvenska skogarnas regionala historia under postarktisk tid. Geologiska Föreningens Förhandlingar 46:83–128

Williams JW, Shuman BN, Webb T, Bartlein PJ, Leduc PL (2004) Late-quaternary vegetation dynamics in north america: Scaling from taxa to biomes. Ecol Monogr 74 (2):309-334. doi:10.1890/02-4045

Wohlfarth B, Skog G, Possnert G, Holmquist B (1998) Pitfalls in the AMS radiocarbon-dating of terrestrial macrofossils. Journal of Quaternary Science 13 (2):137-145